



Extreme Value Theory and Value at Risk: Application to Oil Market

Vêlayoudom Marimoutou, Bechir Raggad, Abdelwahed Trabelsi

► To cite this version:

Vêlayoudom Marimoutou, Bechir Raggad, Abdelwahed Trabelsi. Extreme Value Theory and Value at Risk: Application to Oil Market. 2006. halshs-00410746

HAL Id: halshs-00410746

<https://shs.hal.science/halshs-00410746>

Preprint submitted on 24 Aug 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

GREQAM

Groupement de Recherche en Economie
Quantitative d'Aix-Marseille - UMR-CNRS 6579
Ecole des Hautes Etudes en Sciences Sociales
Universités d'Aix-Marseille II et III

**Document de Travail
n°2006-38**

EXTREME VALUE THEORY AND VALUE AT RISK: APPLICATION TO OIL MARKET

**Vêlayoudom MARIMOUTOU
Bechir RAGGAD
Abdelwahed TRABELSI**

July 2006

DT-GREQAM

Extreme Value Theory and Value at Risk: Application to Oil Market

Velayoudom Marimoutou^a, Bechir Raggad^{b,*}, Abdelwahed Trabelsi^b

^a GREQAM, Université de la Méditerranée, France.

^b BESTMOD, Institut Supérieur de Gestion de Tunis, Tunisie.

July 18, 2006

Abstract

Recent increases in energy prices, especially oil prices, have become a principal concern for consumers, corporations, and governments. Most analysts believe that oil price fluctuations have considerable consequences on economic activity. Oil markets have become relatively free, resulting in a high degree of oil-price volatility and generating radical changes to world energy and oil industries. As a result oil markets are naturally vulnerable to significant negative volatility. An example of such a case is the oil embargo crisis of 1973. In this newly created climate, protection against market risk has become a necessity. Value at Risk (VaR) measures risk exposure at a given probability level and is very important for risk management. Appealing aspects of Extreme Value Theory (EVT) have made convincing arguments for its use in managing energy price risks. In this paper, we apply both unconditional and conditional EVT models to forecast Value at Risk. These models are compared to the performances of other well-known modelling techniques, such as GARCH, historical simulation and Filtered Historical Simulation. Both conditional EVT and Filtered Historical Simulation procedures offer a major improvement over the parametric methods. Furthermore, GARCH(1, 1)-t model may provide equally good results, as well as the combining of the two procedures.

Keywords: Extreme Value Theory, Value at Risk, oil price volatility, GARCH, Historical Simulation, Filtered Historical Simulation.

JEL Classification: C22, C52, G13, Q40.

*Corresponding author, Tel. : +216 98 663 198.

E-mail addresses : raggadbechir@yahoo.fr (Raggad) ; marimout@ehess.univ-mrs.fr (Marimoutou)

1. Introduction

Recent increases in energy prices, especially oil prices, have become a principal concern for consumers, corporations, and governments. Oil as a primary source of energy is needed for industrial production, electric power generation, and transportation. The major oil price shocks during the last three decades were the 1973 oil embargo, the 1979-80 events in Iran and Iraq, the 1990 invasion of Kuwait and latest crisis in the crude oil market in 1999-2000. Numerous studies were carried out to investigate possible effects of oil price fluctuations on the main economic indicators of oil-importing countries. Most analysts believe that oil price fluctuations have considerable consequences on economic activity. Among others, Sadorsky (1999) suggests changes in oil prices have an impact on economic activity, but changes in economic activity have little impact on oil prices. He shows that oil price volatility shocks have asymmetric effects on the economy and finds evidence of the importance of oil price movements when explaining movements in stock returns.

Oil prices were primarily determined by long-term contracts between oil producers and international oil companies. OPEC¹ the most oil dominate the price and quantity of oil sold. Prices fluctuated when these long-term contracts were revised, but prices were not otherwise particularly responsive to market conditions. However, the oil market began to change and become relatively free, resulting in a high degree of oil-price volatility and generating radical changes to world energy and oil industries. Oil market is therefore naturally vulnerable to volatility (see Fattouh, 2005) as it's very linked to the sources and the potentially sources of instability, including political and economic factors². In this newly created climate and in response to an unpredictable, volatile and risky environment, protection against market risk has become a necessity. It is therefore important to model these oil price fluctuations and implement an effective tool for energy price risk management. Value at Risk has become a popular risk measure in the financial industry, whose origins date back to the early 1990's at J.P. Morgan. VaR answers the question about how much we can lose with a given probability over a certain time horizon. The great popularity that

¹ Organization of Petroleum Exporting Countries: the cartel controls 70 percent of the world's known oil reserves and contributes to about 40 percent to world oil production.

² For a discussion of the articulation between economic and political factors in the formation of petroleum prices (see Giraud 1995).

this instrument has achieved is essentially due to its conceptual simplicity: VaR reduces the (market) risk associated with any portfolio to just one number, the loss associated to a given probability.

Existing approaches for VaR estimation may be classified into three approaches. First, the non-parametric historical simulation (HS) approach provides simple empirical quantiles based on the available past data. Second, fully parametric models approach based on an econometric model for volatility dynamics and the assumption of conditional normality (J.P. Morgan's Riskmetrics and most models from the GARCH family) describe the entire distribution of returns including possible volatility dynamic. Third, extreme value theory approach parametrically models only the tails of the return distribution. Since VaR estimations are only related to the tails of a probability distribution, techniques from EVT may be particularly effective. Appealing aspects of EVT have made convincing arguments for its use in calculating VaR and for risk management in general.

Extreme value theory has been applied in many fields where extreme values may appear. Such fields range from hydrology (Davison and Smith, 1990; Katz et al.; 2002) to insurance (McNeil, 1997; Rootzen and Tajvidi, 1997) and finance (Danielsson and de Vries, 1997; McNeil, 1998; Embrechts, 1999; Gençay and Selçuk, 2004). EVT provides a solid framework to formally study the behavior of extreme observations. It focuses directly on the tails of the sample distribution and could, therefore, potentially perform better than other approaches in terms of predicting unexpected extreme changes (see for exemple, Dacorogna et al., 1995; Longin, 2000). However, none of these studies has reflected the current volatility background. In order to overcome this shortcoming, McNeil and Frey (2000) proposed a combined approach that reflects two stylized facts exhibited by most financial return series, namely stochastic volatility and the fat-tailedness of conditional return distributions (see, Pagan, 1996).

Within oil markets, implementing a risk measurement methodology based on the statistical theory of extremes is an important issue. To the best of our knowledge, very few studies have focused on measuring the risk forecasts in the oil market despite the significant need and interest to manage energy price risks. Among the few studies on estimating VaR on energy market with EVT, is the paper of Krehbiel and Adkins (2005) who analyzed the price risk in the NYMEX Energy Complex using an extreme value theory

approach. Given the importance for an effective price risk management tool, a more comprehensive study seems prudent.

The purpose of this paper is to investigate the relative predictive performance of various alternative Value at Risk (VaR) models. Our main focus is on extreme value theory-based models. To this end, both unconditional and conditional Extreme Value Theory (EVT) models are used to forecast VaR. These models are compared to the performances of other well-known modeling techniques, such as GARCH, historical simulation, and Filtered Historical Simulation. Backtesting criteria (unconditional and conditional coverage) are implemented to test the statistical accuracy of the candidate models. Results show that Conditional EVT and Filtered Historical Simulation procedures offer a major improvement over the parametric methods. Furthermore, GARCH (1,1)-t model may give equally good results, as well as the combining of the two procedures.

This paper is organized as follows: Section 2 briefly reviews the principles of risk measurement and extreme value theory. Section 3 discusses various parametric and non-parametric methods that we apply in order to forecast risk measures. Section 4 discusses threshold choice for EVT. Section 5 presents the evaluating framework of VaR models. Section 6 provides our empirical results and Section 7 concludes the paper.

2 Risk measures and Extreme Value Theory

2.1 Risk measures

Value at Risk (VaR) is a popular risk measure in the financial industry, whose origins date back to the early 1990's at J.P. Morgan. VaR is defined as the maximum loss that will be incurred on the portfolio with a given level of confidence over a specified period. In other words, for a given time horizon t and confidence level q , the VaR is the loss in market value over the time t that is exceeded with probability $1-q$. For example, if q is equal to 99% and the holding period is one day, the actual losses on portfolio should exceed VaR estimate not more than once in 100 days on average.

From a statistical point of view, VaR entails the estimation of a quantile of the distribution of returns. To define VaR precisely, let X be the random variable whose cumulative distribution function F_X describes the negative³ profit and loss distribution (P&L) of the risky financial position at the specified horizon time. Formally, value at

³ Negative values of X correspond now to profits and positive values of X correspond to losses.

risk is a quantile of the probability distribution F_X , that is roughly, the x corresponding to a given value of $0 < q = F_X(x) < 1$.

$$\text{VaR}_q(X) = F^{-1}(q), \quad (1)$$

where F^{-1} denotes the inverse function of F_X .

Artzner et al. (1997, 1999) show that VaR has various theoretical deficiencies as a measure of market risk. They conclude that the VaR is not a coherent measure of risk as it fails to be subadditive in general. On the other hand, VaR gives only a lower limit of the losses that occur with a given frequency, but tell us nothing about the potential size of the loss given that a loss exceeding this lower bound has occurred. These authors propose the use of the so-called expected shortfall or tail conditional expectation instead. The expected shortfall measures the expected loss given that the loss L exceeds VaR. In particular, this risk measure gives some information about the size of the potential losses given that a loss bigger than VaR has occurred. Expected shortfall is a coherent⁴ measure of risk as defined by Artzner et al. (1999). Formally, the expected shortfall for risk X and high confidence level q is defined as follows:

$$S_q(X) = E(X | X > \text{VaR}_q(X)) \quad (2)$$

2.2 Extreme Value Theory

The purpose of this section is to summarize the set of results of the extreme value theory necessary to develop the theoretical foundation for this paper. Readers interested in a more detailed background may consult various texts on EVT such as Embrechts et al. (1997) and Reiss and Thomas (1997). The modelling of extremes may be done in two different ways: modelling the maximum⁵ of a collection of random variables, and modelling the largest values over some high threshold. Consequently, we have two significant results: First, the asymptotic distribution of a series of maxima (minima) is modelled and under certain conditions the distribution of the standardized maximum by the generalized extreme value (GEV) distribution. The second significant result concerns the distribution of excess over a given threshold and shows that the limiting distribution is a generalized Pareto distribution (GPD).

2.2.1 The Generalized Extreme Value

Suppose we have an independent and identically (i.i.d.) sequence of random variables X_1, X_2, \dots, X_n representing risks or losses

⁴ A risk measure that satisfies monotonicity, translation invariance, homogeneity and sub-additivity properties is called coherent.

⁵ $\text{Max}(X_1, X_2, \dots, X_n) = \text{Min}(-X_1, -X_2, \dots, -X_n)$ Therefore, all the results for the distribution of maxima leads to an analogous result for the distribution of minima and vice versa. We will discuss the results for maxima only.

with unknown common cumulative distribution function (c.d.f), $F(x) = \Pr(X_i \leq x)$. As a convention, a loss is treated as a positive number and extreme events take place when losses come from the right tail of the loss distribution F . Let $M_n = \max(X_1, X_2, \dots, X_n)$ denote the n^{th} sample maximum in a sample of n losses. We are interested in the behavior of M_n as n approaches infinity. For a sample of i.i.d. observations, the c.d.f of M_n is given by

$$\Pr(M_n \leq x) = F(x)^n \quad (3)$$

This result implies that $\Pr(M_n \leq x)$ approaches 0 or 1 depending whether F has a finite upper end-point or not, as n approaches infinity. Let x_F be the finite or infinite upper end-point of the distribution F such that $x_F = \{x: F(x) < 1\}$.

While result (3) is of no immediate interest, it tell us nothing about the distribution of M_n for large n , we rely on the Fisher-Tippett theorem to examine the asymptotic behavior of the distribution. It does for the maxima of i.i.d. random variables what the central limit theorem does for sums. Fisher-Tippett showed that if there exist norming constants $a_n > 0$ and $b_n \in \mathbb{R}$ and some non-degenerate distribution function H such that: $\frac{M_n - b_n}{a_n} \xrightarrow{d} H$ then H is one of the following three⁶ types:

Weibull:	$\Psi_\alpha(x) = \begin{cases} \exp[-(-x)^\alpha], & x \leq 0, \alpha > 0 \\ 1, & x > 0 \end{cases}$
Gumbel:	$\Lambda(x) = \exp(-e^{-x}), \quad x \in \mathbb{R}$
Fréchet:	$\Phi_\alpha(x) = \begin{cases} 0, & x \leq 0 \\ \exp[-x^{-\alpha}], & x > 0, \alpha > 0 \end{cases}$

By taking the reparameterization $\xi = 1/\alpha$ due to Jenkinson (1955) and von Mises (1936), Weibull, Gumbel and Fréchet distributions can be represented in a unified model with a single parameter. This reparameterized version, H , is called the generalized extreme value (GEV) distribution (see Embrechts et al., 1997, pp.121-152).

⁶ The distributions $\Psi_\alpha(x)$, $\Lambda(x)$ and Φ_α are called standard extreme value distributions.

$$H_{\xi}(x) = \begin{cases} \exp(-(1+\xi x)^{-\frac{1}{\xi}}), & \xi \neq 0, 1+\xi x > 0 \\ \exp(-e^{-x}), & \xi = 0, -\infty \leq x \leq +\infty \end{cases} \quad (4)$$

In this case, X (and the underlying distribution F) is said to belong to the Maximum Domain of Attraction of the extreme value distribution of H and we write $X \in \text{MDA}(H_{\xi})$.

The most important parameter is ξ , which indicates the thickness of the tail of the distribution. The larger the tail index, the thicker the tail. It determines, essentially, the tail behavior of H . Distributions that belong to $\text{MDA}(H_{\xi})$, for $\xi > 0$ are called heavy⁷ tailed (examples are Pareto, Cauchy, Student-t, loggamma). In practice, modeling all block maxima is wasteful if other data on extreme values are available. Therefore, a more efficient approach is to model the behavior of extreme values above a high threshold.

2.2.2 The Generalized Pareto Distribution

Suppose that X_1, X_2, \dots, X_n are n independent realizations of a random variable X with a distribution function $F(x)$. Let x_F be the finite or infinite right endpoint of the distribution F . The distribution function of the excesses X_i over certain high threshold u is given by

$$F_u(y) = \Pr(X - u \leq y \mid X > u) = \frac{F(y+u) - F(u)}{1 - F(u)}, \quad y \geq 0 \quad (5)$$

Balkema and de Haan (1974) and Pickands (1975) theorem⁸ showed that for a certain class of distribution functions, the generalized Pareto distribution (GPD) is the limiting distribution for $F_u(y)$ as the threshold tends to the right endpoint. They stated that if the distribution function $F \in \text{DAM}(H_{\xi})$ then it is possible to find a positive measurable function $\beta(u)$ such that:

$$\lim_{u \rightarrow x_F} \sup_{0 \leq y \leq x_F - u} |F_u(y) - G_{\xi, \beta(u)}(y)| = 0 \quad (6)$$

⁷ Thin-tailed distributions include the normal, exponential, gamma and lognormal belong to $\text{MDA}(H_0)$, while distributions with a finite right-hand end points (such as the uniform and beta distributions) belong to $\text{MDA}(H_{\xi})$, for $\xi < 0$.

⁸ For more details consult Theorem 3.4.13 on page 165 of Embrechts et al. (1997).

where $G_{\xi, \beta(u)}(y)$, the GPD, is

$$G_{\xi, \beta(u)}(y) = \begin{cases} 1 - \left(1 + \xi \frac{y}{\beta(u)}\right)^{-\frac{1}{\xi}}, & \xi \neq 0 \\ 1 - \exp\left(-\frac{y}{\beta(u)}\right), & \xi = 0 \end{cases} \quad (7)$$

where $y \geq 0$ for $\xi \geq 0$ and $0 \leq y \leq -\frac{\beta(u)}{\xi}$ for $\xi < 0$. The choice of the threshold u is crucial for the success of the GPD modelling. The appropriate value is typically chosen by a trade-off between bias and variance (see section 4).

ξ is the important shape parameter of the distribution and is an additional scaling parameter $\beta(u)$. The GPD embeds a number of other distributions. If $\xi > 0$ then G is a reparametrized version of the ordinary Pareto distribution, which has a long history in actuarial mathematics as a model for large losses; $\xi = 0$ corresponds to the exponential distribution and $\xi < 0$ is known as a Pareto type II distribution. The first case is the most relevant for risk management purposes since the GPD is heavy-tailed when $\xi > 0$. Estimates of the parameters ξ and $\beta(u)$ can be obtained from expression (7) by the method of maximum likelihood (see Embrechts et al.(1997)). For $\xi > -0.5$, Hosting and Wallis (1987) present evidence that maximum likelihood regularity conditions are fulfilled and the maximum likelihood estimates are asymptotically normally distributed.

3 Statistical Approaches to Value-at-Risk

In this section, we present the various parametric and non-parametric methods that we use to forecast risk measures, VaR_q , in the oil market. Our main interest is on extreme value theory based modes: we consider the unconditional GPD, the GARCH models, the conditional GPD, the historical simulation, filtered historical Simulation and finally we include the unconditional normal model where we assume that the returns come from the normal distribution with historically estimated mean and variance.

3.1 The Peaks-over-Threshold Model: the GPD approach

Since VaR estimations are only related to the tails of a probability distribution, techniques from EVT may be particularly effective. Appealing aspects of EVT have made convincing arguments for its use in calculating VaR . As we have discussed, the modelling of extremes may be done in two different ways: modelling the maximum of a collection of random variables, and modelling the largest values over some high threshold, known as the 'Peaks-Over-Threshold

(POT)' model. In this paper, we use the latter, more modern approach to modelling extreme events. The POT models are generally considered to be more appropriate for practical applications, due to their more efficient use of the limited data as all observations above the threshold are utilized to estimate parameters of the tail.

Our approach to the GPD modeling is as follows. We fix a sufficiently high threshold u . Let $Y_1 \dots Y_n$ be the excesses above this threshold where $Y_i = X_i - u$. Balkema and de Haan (1974) and Pickands (1975) theorem (6) justify that $F_u(y) = G_{\xi, \beta(u)}(y)$ provided the threshold is sufficiently high. By setting $x = u + y$, an approximation of $F(x)$, for $x > u$, can be obtained from equation (5):

$$F(x) = (1 - F(u))G_{\xi, \beta(u)}(y) + F(u), \quad (8)$$

and the function $F(u)$ can be estimated non-parametrically using the empirical c.d.f:

$$\hat{F}(u) = \frac{n - N_u}{n}, \quad (9)$$

where N_u represents the number of exceedences over the threshold u and n is the sample. After substituting equations (7) and (9) into equation (8), we get the following estimate for $F(x)$

$$\hat{F}(x) = 1 - \frac{N_u}{n} \left(1 + \hat{\xi} \left(\frac{x - u}{\hat{\beta}} \right) \right)^{-\frac{1}{\hat{\xi}}}, \quad (10)$$

where $\hat{\xi}$ and $\hat{\beta}$ are estimates of ξ and β , respectively, which can be obtained by the method of maximum likelihood.

For $q > F(u)$, VaR_q can be obtained from (10) by solving for x

$$\text{VaR}_q = \hat{u} + \frac{\hat{\beta}}{\hat{\xi}} \left[\left(\frac{n}{N_u} (1 - q) \right)^{-\hat{\xi}} - 1 \right] \quad (11)$$

where u is a threshold, β is the estimated scale parameter, ξ is the estimated shape parameter.

The main advantage of unconditional GPD approach is that it focuses attention directly on the tail of the distribution. However, it

doesn't recognize the fact that returns are non-i.i.d

3.2 GARCH Modelling

The issue of modelling returns that account for time-varying volatility has been widely analyzed in financial econometrics literature. Two main types of techniques have been used: Generalized Autoregressive Conditional Heteroscedasticity (GARCH)-models Bollerslev (1986) and stochastic volatility models. GARCH processes have gained fast acceptance and popularity in the literature devoted to the analysis of financial time series. These time series models capture several important features of the financial series, such as volatility clustering and leptokurticity.

As opposed to the EVT-based models described above, GARCH models do not focus directly on the returns in the tails. Instead, by acknowledging the tendency of financial return volatilities to be time-dependent, GARCH models explicitly model the conditional volatility as a function of past conditional volatilities and returns. Let $\{X_t\}_{t=0}^n$ refer to be the negative return series. We assume that the dynamics of the return series follows the stochastic process:

$$\begin{aligned} X_t &= \mu_t + \varepsilon_t = \mu_t + z_t \sigma_t \\ z_t &\stackrel{\text{i.i.d}}{\sim} f(E(z_t) = 0, V(z_t) = 1) \\ \sigma_t^2 &= g(\Omega_{t-1}), \end{aligned} \quad (12)$$

where $E(X_t | \Omega_{t-1}) = \mu_t$ denotes the conditional mean, given the information set available at time $t-1$, Ω_{t-1} , $\{\varepsilon_t\}_{t=0}^n$ is the innovation process with conditional variance $V(X_t | \Omega_{t-1}) = \sigma_t^2$, $f(\cdot)$ is the density function of $\{z_t\}_{t=0}^n$ and g is the functional form of the conditional volatility model. In estimating VaR with GARCH type models it 'is commonly supposed that the innovation distribution follows a normal distribution "conditional normal distribution" so that an estimate of VaR is given by the following equation:

$$\text{VaR}_{t+1, q} = \mu_{t+1} + \sigma_{t+1} \Phi^{-1}(q), \quad (13)$$

where $\Phi^{-1}(\cdot)$ is the quantile of the standard normal distribution, μ_{t+1} and σ_{t+1} are the conditional forecasts of the mean and the standard deviation at time $t+1$, given the information at time t . It is generally well recognized that GARCH-models coupled with conditionally normally distributed innovations "conditional student distribution" is unable to fully account for the tails of the marginal distributions of daily returns. Several alternative conditional distributions have therefore been proposed in the GARCH (e.g. Student-t distribution, generalized error distribution (GED), and etc.).

In this paper, we will show that GARCH models with Student's t-distribution yields quite satisfactory results. In this case, the VaR is given by:

$$\text{VaR}_{t+1, q} = \mu_{t+1} + \sigma_{t+1} \sqrt{\frac{v-2}{v}} F^{-1}(q), \quad (14)$$

where $F^{-1}(q)$ is the quantile the t-distribution with v degrees of freedom ($v > 2$).

GARCH modeling approach that does model the conditional return distribution as time varying, but focuses on the whole return distribution and not only on the part we are primarily interested in, the tail. Therefore, this approach may fails to estimate accurately the risk measures like VaR.

3.3 Conditional GPD

In order to overcome the drawbacks of both approaches presented above, the immediate solution is to combine these two approaches as firstly suggested by McNeil and Frey (2000). The combined approach, denoted conditional GPD, has the following steps:

- Step 1: Fit a GARCH-type model to the return data by quasi-maximum likelihood. That is, maximize the log-likelihood function of the sample assuming normal innovations. Estimate μ_{t+1} and σ_{t+1} from the fitted model and extract the residuals z_t .
- Step 2: Consider the standardized residuals computed in Step 1 to be realizations of a white noise process, and estimate the tails of the innovations using extreme value theory. Next, compute the quantiles of the innovations for $q = 0.95$.
- Step 3: Construct VaR (ES) from parameters estimated in steps one and two.

We assume that the dynamics of log-negative returns can be represented by (12). Given the 1-step forecasts μ_{t+1} , σ_{t+1} and the estimate quantile of standardized residuals series, $\text{VaR}_{t+1}(Z)$, using the equation (11) the VaR for the return series can be estimated as:

$$\text{VaR}_{t+1, q} = \mu_{t+1} + \sigma_{t+1} \text{VaR}_t(Z), \quad (15)$$

By first filtering the returns with a GARCH model is that we get essentially i.i.d. series on which it is straightforward to apply the EVT technique. The main advantage of the conditional GPD is that it produces a VaR which reflect the current volatility background.

3.4 Historical Simulation

The first and the most commonly used method is referred to as the historical simulation (HS) approach. The main idea behind the HS is the assumption that historical distribution of returns will remain the same over the next periods: it assumes that price change behavior repeats itself over time. Therefore, future distribution of returns is well described by the empirical, historical return that will be used in estimating VaR. As a result, the VaR based on HS is simply the empirical quantile of this distribution associated with the desired likelihood level.

$$\text{VaR}_{t+1, q} = \text{Quantile} \{ \{X_t\}_{t=1}^n \}, \quad (16)$$

Historical Simulation (HS) has a number of advantages. It is easy to understand and to implement. It's completely nonparametric and does not depend on any distribution assumption, thus capturing the non-normality in the data. HS also has several disadvantages. Most notably, it is impossible to obtain an out-of-sample VaR estimate with HS. HS ignores the potentially useful information in the volatility dynamics. For a complete discussion on the use of historical simulation approach for VaR estimation, you can see various articles such as Hendricks (1996) and Barone et al. (2000).

3.5 Filtered Historical Simulation

In order to remedy some of the shortcomings of the simulation approach, we apply the filtered historical simulation (FHS) approach introduced by Hull and White (1998) and Barone-Adesi et al. (1999). This approach combines the historical simulation and the GARCH models. Specifically, it does not make any distributional assumption about the standardized returns, while it forecasts the variance through a volatility model. Hence, it is a mixture of parametric and non-parametric statistical procedures. Moreover, Barone-Adesi and Giannopoulos (2001) demonstrated the superiority of the filtered historical simulation over the historical one, since it generated better VaR forecasts than the latter method. The main advantage of FHS is that it can produce risk measures that are consistent with the current state of markets at any arbitrarily large confidence level.

FHS consists on fitting a GARCH-model to return series and use historical simulation to infer the distribution of the residuals. By using the quantiles of the standardized residuals, the conditional standard deviation and the conditional mean forecasts from a volatility model, the VaR number is given as:

$$\text{VaR}_{t+1, q} = \mu_{t+1} + \sigma_{t+1} \text{Quantile} \{ \{X_t\}_{t=1}^n \}, \quad (17)$$

The combination of the two methods might lessen the problematic use of the “classical” approaches, since this procedure can accommodate the volatility clustering, the observed “fat” tails and the skewness of the empirical distribution.

4 Threshold choice for EVT

Balkema and de Haan (1974) and Pickands (1975) theorem (6), tells us that above sufficiently high thresholds the distribution of the excesses may be approximated by a GPD. The parameters of the GPD may be estimated by using, for example, maximum likelihood once the threshold u has been chosen. However, this choice is subject to a trade-off between variance and bias. By increasing the number of observations for the series of maxima (a lower threshold), some observations from the centre of the distribution are introduced in the series, and the index of the tail is more precise but biased (i.e., there is less variance). On the other hand, choosing a high threshold reduces the bias but makes the estimator more volatile (i.e., there are fewer observations). The problem of finding an optimal threshold is very subjective: we need to find a threshold u above which the Pickands, Balkema and de Hann theorem (3.4) holds and the GDP is a reasonable model of exceedances. However, the threshold must also be chosen such that we have sufficient data to accurately estimate parameters of the distribution.

There is no unique choice of the threshold level. A number of diagnostic techniques exist for this purpose, including graphical, bootstrap methods [see Embrechts et al. (1997), Reiss and Thomas (1997)].

In this paper, a simulation exercise is conducted. We generated samples of size $n=1000$, using two different distributions, the i.i.d. symmetric student- $t(\nu)$ with $\nu = 4,6$ and the GARCH(1,1) with student- $t(\nu)$ innovations. We take the parameterization used by Wagner and Marsh (2004). The distributions are all in the maximum domain of attraction of the Fréchet with ξ parameter 0.25 or 0.17. This particular choice is driven by two main motivations. As we will show later, the student- t may provide a rough approximation to the observed distribution of model residuals. On the other hand, it allow us to compare the dependant GARCH (1, 1)- t models to the i.i.d student- t .

We chose the threshold values indirectly, by choosing the k number of exceedances (k) to be included in the maximum likelihood

estimation. We started with $k = 20$ and we increased it by 1 until it reached 500. To compare the different estimates, we computed the bias and the mean squared error of the estimators as follows; the bias of the estimator is defined to be $\text{Bias}(\hat{\xi}_k) = E[\hat{\xi}_k - \xi]$; the expected difference between the estimator and the true tail index value; it is estimated in our study by

$$\text{Bias} = \frac{1}{100} \sum_{i=1}^{100} \xi_k^{(i)} - \xi, \quad (18)$$

where $\xi_k^{(i)}$ denotes the i th MLE estimate of obtained from the i th sample. The mean square error of the tail index estimator is defined to be $\text{MSE}(\hat{\xi}_k) = E(\hat{\xi}_k - \xi)^2$, and it can be shown that $\text{MSE}(\hat{\xi}_k) = \text{Var}(\hat{\xi}_k) + \text{Bias}(\hat{\xi}_k)^2$. In our study we estimate the $\text{MSE}(\hat{\xi}_k)$ by

$$\text{MSE}(\hat{\xi}_k) = \frac{1}{100} \sum_{i=1}^{100} (\xi_k^{(i)} - \xi)^2, \quad (19)$$

Our objective is to determine how sensitive these estimates are to the choice of the parameter k , to the underlying distribution. In Figure 1, we plot the bias and MSE of the EVT estimates against k . Graphical inspection of this figure shows that a value of k between 100 and 150 may be justified for the two distributions. We choose⁹ $k=140$ for our GPD approach.

5 Evaluating VaR models

We test the reliability of our VaR methodology by investigating the out-of sample performance of the estimated VaRs in forecasting extreme returns. The backtesting procedure consists of comparing the VaR estimates with actual realized loss in the next period. Two backtesting criteria¹⁰ are implemented for examining the statistical accuracy of the models. First, we determine whether the frequency of exceedances is in line with the predicted confidence level VaR based on the unconditional coverage test of Kupiec (1995). However, tests of unconditional coverage fail to detect violations of independence property of an accurate VaR measure, it's important to examine if the violations are also randomly distributed. Second, given that an accurate VaR model must exhibit both the unconditional coverage and independence property, we test jointly both properties based on conditional coverage Test of Christoffersen (1998).

⁹ In a similar simulation exercise McNeil and Frey (2000) concluded, in the iid case, that a k value of 100 seems reasonable, but they argued that could equally choose a k value of 80 or 150.

¹⁰ For a review of different backtesting procedures see (Campbell 2005).

5.1 Unconditional Coverage Test

Let $N = \sum_{t=1}^T I_{t+1}$ be the number of days over a T period that the portfolio loss was larger than the VaR estimate, where I_{t+1} be a sequence of VaR violations¹¹ that can be described as:

$$I_{t+1} = \begin{cases} 1, & \text{if } X_{t+1} < \text{VaR}_{t+1} \mid t \\ 0, & \text{if } X_{t+1} \geq \text{VaR}_{t+1} \mid t. \end{cases}$$

We use a likelihood ratio test developed by Kupiec (1995). This test examines whether the failure rate is statistically equal to the expected one. Let p be the expected failure rate ($p = 1 - q$, where q is the confidence level for the VaR). If the total number of such trials is T , then the number of failures N can be modelled with a binomial distribution¹² with probability of occurrence equals to p . The correct null and alternative hypothesis are, respectively $H_0: \frac{N}{T} = p$ and $H_1: \frac{N}{T} \neq p$.

The appropriate likelihood ratio statistic is:

$$LR_{uc} = 2 \left[\log \left(\left(\frac{N}{T} \right)^N \left(1 - \frac{N}{T} \right)^{T-N} \right) - \log(p^N (1 - p)^{T-N}) \right] \quad (20)$$

$LR_{uc} \xrightarrow{d} \chi^2(1)$ under H_0 of good specification. Note that this backtesting procedure is a two sided test. Therefore, a model is rejected if it generates too many or too few violations, but based on it, the risk manager can accept a model that generates dependent exceptions. Accordingly, this test may fail to detect VaR measures that exhibit correct unconditional coverage but exhibit dependent VaR violations. So we turn to a more elaborate criterion.

5.2 Conditional Coverage Test

Christofersen (1998) proposed a more comprehensive and elaborate test, which jointly investigates if (i) the total number of failures is equal to the expected one and (ii) the VaR failure process is independently distributed through time. This test provides an opportunity to detect VaR measures which are deficient in one way or the other. Under the null hypothesis that the failure process is

¹¹ A violation occurs if the forecasted VaR is not able to cover realized loss.

¹² $X \sim B(n, p) : P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}$

independent and the expected proportion of violations is equal to p , the appropriate likelihood ratio is:

$$LR_{cc} = -2\log[(1-p)^{T-N} p^N] + 2\log[(1-\pi_{01})^{n_{00}} \pi_{01}^{n_{01}} (1-\pi_{11})^{n_{10}} \pi_{11}^{n_{11}}] \xrightarrow{d} \chi^2(2), \quad (21)$$

where n_{ij} is the number of observations with value i followed by j , for $i, j = 0, 1$ and $\pi_{ij} = \frac{n_{ij}}{\sum_j n_{ij}}$ are the corresponding probabilities. The

values $i, j = 1$ denote that a violation has been made, while $i, j = 0$ indicates the opposite. The main advantage of this test is that it can reject a VaR model that generates either too many or too few clustered violations.

6 Empirical analysis

6.1 Data

The data for our analysis consists of the daily spot Brent oil prices, over the period May 21, 1987 through January 24, 2006 excluding holidays. By using this time period, we get a complete sample containing 4810 observations. Figure 2 shows oil price trends corresponding to the analyzed period and indicates that the oil price has mainly fluctuated in the range of about 9-67 dollars.

[Insert Figure 2 about here]

The sample mean and standard deviation of the oil price in this period are about 18 and 4 dollars, respectively. From these prices we calculate 4809 log-returns and plot them in figure 3. This graphic show that returns are stationary and suggests an ARCH scheme for the daily oil price returns where large changes are followed by large changes and small changes are followed by small changes.

[Insert Figure 3 and Table 1 about here]

Table 1 provides a summary statistics on the return series, the Jarque-Bera statistic shows that the null hypothesis of normality is rejected at any level of significance, as evidenced by high excess kurtosis and negative skewness. The unconditional distribution is non-normal has a long left tail relative to a symmetric distribution. The Ljung-Box statistic for serial correlation shows that null hypothesis of no autocorrelation for up to 20th order is rejected at any level of significance and confirms the presence of conditional heteroskedasticity. It is also important to note that the returns series are inconsistent with the necessary condition of the extreme value theory, i.e. that samples are independent and identically distributed. To overcome this shortcoming, it is necessary to filter returns with a GARCH model in order to get essentially i.i.d. series on which it is straightforward to apply the EVT.

6.2 Modeling oil price volatility

The result of a specification search in terms of AIC and BIC criteria for a wide range of values for p and q leads us to choose the ar(1)-garch(1,1) model, given by the following equation, as the best model:

$$X_t = \mu_t + Z_t \sigma_t, \quad (22)$$

where Z_t are i.i.d innovations with zero mean and unit variance, and

$$\mu_t = \alpha_0 + \alpha_1 X_{t-1}$$

$$\sigma_t^2 = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \gamma \sigma_{t-1}^2$$

where μ_t and σ_t^2 denote the conditional mean and the conditional variance of the process. This model is fitted to data series using a pseudo-maximum likelihood estimation assuming normal distributed

innovations to obtain parameter estimates $\hat{\theta} = \left(\hat{\alpha}_0, \hat{\alpha}_1, \hat{\beta}_0, \hat{\beta}_1, \hat{\gamma} \right)$ and

standardized residuals $\frac{(X_t - \mu_t)}{\sigma_t}$.

[Insert Table 2 about here]

Table 2 presents the estimated parameters of the mean and volatility equations of oil returns. Both the constant term and ar(1) coefficient in the mean equation are found to be significant. Similarly, the parameters in the volatility equations: the constant, the arch(1) parameter and the garch (1) parameter, are all found to be significant.

In this paper, the AR(1) GARCH(1,1) model is used in three contexts. In the first context, AR(1)-GARCH(1,1)-normal, the model is used directly as a risk measurement methodology for comparison with other candidate risk measurement methods. In the second context, the conditional GPD, the model serves to pre-filter the data series and produce the standardized residuals used to estimate tail parameters with the POT methodology. Finally, Filtered historical simulation, the model serves again to pre-filter the data series and produce the standardized residuals which will be inferred by the HS approach.

We estimate the AR (1) GARCH (1, 1) specification using a rolling window of 1000 days data. We extract the standard residuals from the estimated model for two reasons: (i).to investigate the adequacy of ARCH modelling, and (ii), to use in the combined approaches described above (Conditional GPD, FSH). Table 3 and

figure 4 illustrate the effect of filtering the raw data with a GARCH modelling. Results indicate that the return series have significant ARCH effects, excess kurtosis and autocorrelation. The residual series is found to have significant excess kurtosis but it does possess neither significant autocorrelation nor any ARCH effect left. The results can be summarized in the followings: Neither the return series nor the residual series can be considered to be normally distributed, since both the series have significant excess kurtosis. Therefore, the assumption of conditional normality is unrealistic. On the other hand, the residual series is found to be free from autocorrelation. We can concludes that the filtering process successfully remove the time series dynamics from the return series and obtain an i.i.d series free from any time series dynamics. Therefore, EVT methods may be applied successfully to the i.i.d residual series.

[Insert Figure 4 and Table 3 about here]

6.3 Dynamic backtesting

For all models, we use a rolling sample of 1000 observations, in order to forecast the VaR_q for $q \in \{0.95, 0.99, 0.995, 0.999\}$. The main advantage of this rolling window technique is that it allows us to capture dynamic time-varying characteristics of the data in different time periods. As documented by McNeil and Frey (2000) and Gençay et al. (2003), within the backtest period, it is practically impossible to examine the fitted model carefully every day and to choose the best parameterization, so suppose that the AR(1)GARCH(1,1) specification is adequate on each rolling window. A similar constraint is also related to the GPD modelling. For this reason, we always set $k = 140$ in this backtest, a choice that is supported by the simulation study of section 4.

At each iteration, we compare the predicted VaR number with the realized return, to determine whether the frequency of exceedances is in line with the predicted confidence level of the VaR. If the number of violations is significantly different from the predicted level of violations, then the VaR estimation approach is not valid. Statistically, we use the two backtesting tests (unconditional and conditional coverage tests) explained above to access the statistical accuracy of the various risk management models. The number of violations for various confidence levels and the p-values of the corresponding backtesting measures test are presented in tables 4, 5 and 6. A p-value less than or equal to 0.05 will be interpreted as evidence against the null hypothesis.

[Insert Tables 4, 5 and 6 about here]

The general observation would be that for the 95% VaR measures the EVT-based models and the others traditional models produce equally good VaR estimates (except for the Normal method at the 95% confidence level). As expected, the unconditional normal distribution performs poorly and is rejected for all confidence levels. This model underestimates the “true” VaR and is not appropriate for extreme quantiles estimation. The conditional normal approach can not be rejected for the 95% confidence level but its performance deteriorates at higher quantiles. This approach, while it responds to changing volatility, tends to be violated rather more often, because it fails to fully account to the leptokurtosis of the residuals. This model tends to underestimate the true risk. Such result constitutes an alarm to any market participants that use the models based on normality assumption. Conditional GPD model yields a better VaR estimation than provided by the GDP. The number of days when VaR is higher than actual price change is close to the expected one. Furthermore, Conditional GPD methodology provides a more flexible VaR quantification, which accounts of volatility dynamics.

[Insert Figures 5 and 6 about here]

Figures 5 and 7 show a part of the backtest for oil returns. In figure 5, we have plotted the negative returns; superimposed on this figure is the 99% conditional GPD VaR estimate, the 99% conditional normal VaR estimate and the 99% unconditional GPD VaR estimate. The violations corresponding to the backtest in figure 5 are shown in figure 6. We use different plotting symbols to show violations of the conditional GPD, conditional normal and unconditional GPD quantile estimates. Clearly, the conditional normal estimate responds to volatility dynamics but tends to be violated rather more often as it fails to describe the extreme tails. Conditional Student model performs better than the conditional normal model and provides a very satisfying result, which is very comparable to the conditional GPD. As noted by McNeil and Frey (2000) this method can be viewed as a special case of the conditional GPD approach. It yields quite satisfactory results as long as the positive and the negative tail of the return distribution are roughly equal. Filtered Historical simulation approach is well suited for a VaR estimation and provides an improvement to the standard Historical approach. The FSH is almost close to the mark in VaR estimation. The violations number are too close to the theoretical ones.

[Insert Figures 7 and 8 about here]

In figure 7, we have plotted the negative returns. Superimposed on this figure is the 99% conditional GPD VaR estimate, the 99% conditional Student VaR estimate and the 99% filtered historical simulation VaR estimate. The violations corresponding to the backtest in figure 7 are shown in figure 8. This latter shows that there is not a particular difference between the three models; the violations points are more or less the same.

7 Conclusion

As the volatility in the oil markets increases, implementing an effective risk management system becomes an urgent necessity. In risk management, the VaR methodology as a measure of market risk has gained fast acceptance and popularity in both institutions and regulators. Furthermore, extreme value theory has been successfully applied in many fields where extreme values may appear. VaR methodology benefits from the quality of quantile forecasts. In this paper, mainly EVT models are compared to conventional models such as GARCH, historical simulation and filtered historical. Our results indicate that that Conditional Extreme Value Theory and Filtered Historical Simulation procedures offer a major improvement over the traditional methods. Such models produce a VaR which reacts to the change of volatility dynamics. Furthermore, the GARCH (1, 1)-t model may give equally good results, as well as the two combined approach. Oil price fluctuations are closely linked to economic indicators. For further study, we suggest to study the dependence relation via copula functions.

Acknowledgements

The authors would like to thank Badih Gattas, Claude Deniau and Mhamed ali El Aroui for their helpful comments.

References

- Artzner, P., Delbaen, F., Eber, J., Heath, D., 1997. Thinking coherently. *Risk* 10, 68-71.
- Artzner, P., Delbaen, F., Eber, J., Heath, D., 1999. Coherent measures of risk. *Mathematical Finance* 9 (3), 203-228.
- Barone-Adesi, G., Giannopoulos, K. and Vosper, L., 2000. Backtesting the filtered Historical Simulation. unpublished manuscript.
- Barone-Adesi, G., Giannopoulos, K. and Vosper, L., 1999. VaR without correlations for nonlinear portfolios. *Journal of Futures Markets* 19, 583- 602.
- Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31, 307-327.
- Burns, P., 2002. The quality of value at risk via univariate GARCH. patrick@burns-stat.com<http://www.burns-stat.com>.
- Campbell, S. D., 2005. A review of backtesting and backtesting procedures. Finance and Economics Discussion Series Divisions of Research Statistics and Monetary Affairs Federal Reserve Board, Washington, D.C.
- Christoffersen P., 1998. Evaluating Interval Forecasts. *International Economic Review* 39, 841-862.
- Dacorogna, M. M., Müller, U. A., Pictet, O. V., deVries, C. G., 1995. The distribution of extremal foreign exchange rate returns in extremely large data sets. Preprint, O&A Research Group.
- Danielsson, J., de Vries, C. G., 1997. Tail index estimation with very high frequency data. *Journal of empirical Finance* 4, 241 -257.
- Davison, A. C., Smith, R. L., 1990. Models for exceedances over high thresholds. *Journal of Royal Statistic Society Ser B* 52 (3), 393-442.
- Embrechts, P., Kluppelberg, C., Mikosch, T., 1997. *Modelling Extremal Events for Insurance and Finance*. Springer, Berlin.
- Embrechts, P., Resnick, S., Samorodnitsky, G., 1999. Extreme value theory as a risk management tool. *North Amer. Actuar. J.* 26, 30-41.
- Fattouh, B., 2005. The causes of crude oil price volatility. *Middle East Economic Survey XLVIII* (13-28).
- Gencay, R., Selcuk, F. and Ulugulyagci, A., 2003. High volatility, thick tails and extreme value theory in value-at-risk estimation. *Insurance: Mathematics and Economics* 33, 337-356.
- Gençay, R., Selçuk, F., 2004. Extreme value theory and value-at-risk: Relative performance in emerging markets. *International Journal of Forecasting* 20, 287– 303.
- Giraud, P. N., 1995. The equilibrium price range of oil economics, politics and uncertainty in the formation of oil prices. *Energy Policy* 23 (1), 35-49.
- Hendricks, D., 1996. Evaluation of value-at-risk models using historical data. *Economic Policy Review* 2, 39-70.
- Hosking, J., Wallis, J., 1987. Parameter and quantile estimation for the generalized pareto distribution. *Technometrics* 29, 339- 349.
- Hull, J., White, A., 1998. Incorporating volatility updating for value-at-risk. *Journal of Risk* 1, 5 -19.
- Katz, R. W., Parlange, M. B., Naveau, P., 2002. Statistics of extremes in hydrology. *Advances in Water Resources* 25, 1287-1304.
- Krehbiel, T., Adkins, L. C., 2005. Price risk in the NYMEX energy complex: An extreme value approach. *Journal of Futures Markets* 25(4), 309-337.
- Kupiec P., 1995. Techniques for Verifying the Accuracy of Risk Management Models. *Journal of Derivatives* 3, 73-84.
- Longin, F. M., 2000. From value at risk to stress testing: The extreme value approach. *Journal of Banking and Finance* 24, 1097-1130.
- McNeil, A. J., 1998. Calculating quantile risk measures for financial time series using extreme value theory. Department of Mathematics, ETH. Swiss Federal Technical University E-Collection. <http://e-collection.ethbib.ethz.ch/>.
- McNeil, A. J., Frey, R., 2000. Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *Journal of Empirical Finance* 7, 271-300.
- McNeil, A. J., Saladin, T., 1997. The peaks over thresholds method for estimating high quantiles of loss distributions. Department of Mathematics, ETH Zurich.
- Pagan, A., 1996. The econometrics of financial markets. *Journal of Empirical Finance* 3, 15-102.
- Reiss, R., Thomas, M., 1997. *Statistical Analysis of Extreme Value with Applications to Assurance, Finance, Hydrology and Other Fields*. Birkhauser Verlag, Basel.
- Sadorsky, P., 1999. Oil price shocks and stock market activity. *Energy Economics* 21, 449-469.
- Wagner, N., Marsh, T. A., 2004. Measuring tail thickness under GARCH and an application to extreme exchange rate changes, *Journal of Empirical Finance* 12, 165-185.

Appendix: Figures & Tables

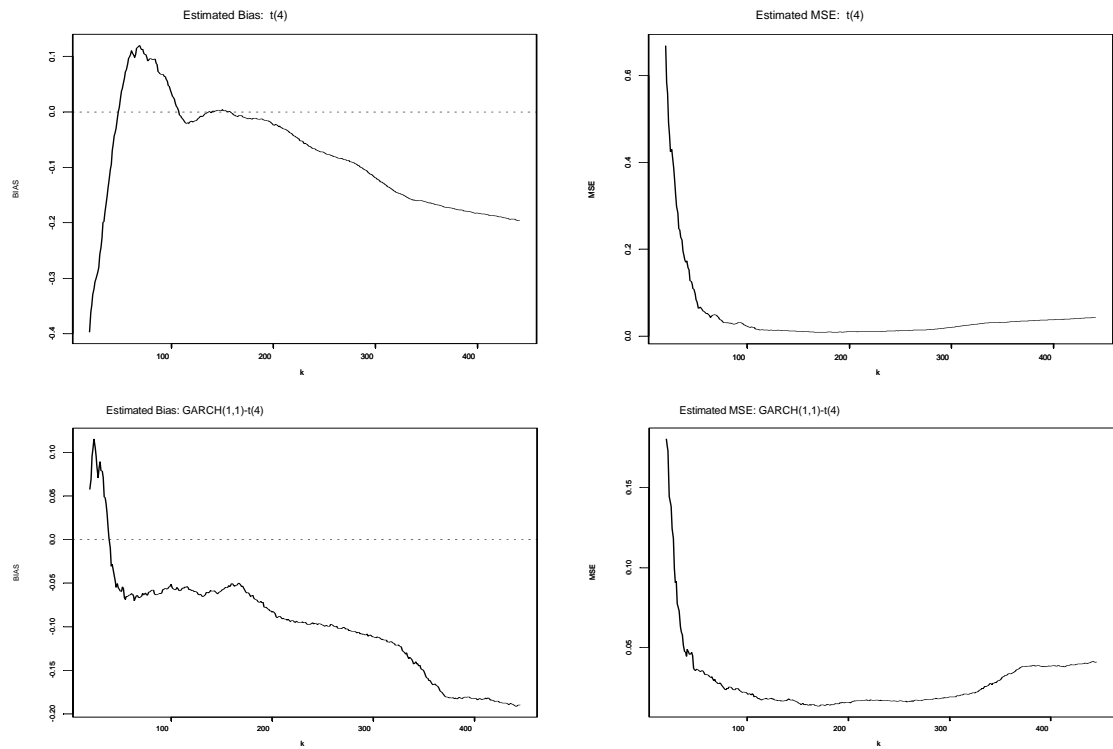


Fig. 1. Estimated bias and Mean Squared Error (MSE) against k for various estimators of shape parameter, ζ_r , of two distributions: a t distribution with $v=4$ degrees of freedom based on an iid sample of 1000 observations and and AR(1)-GARCH(1,1)- $t(v=4)$.

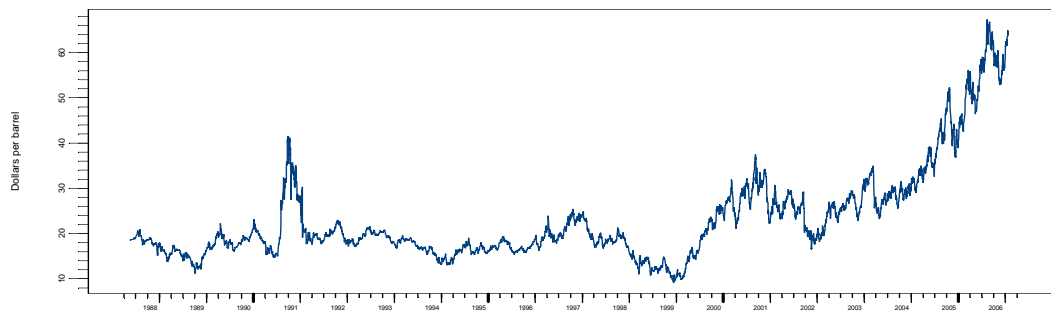


Fig. 2. Daily prices of Brent crude (US \$ per barrel).

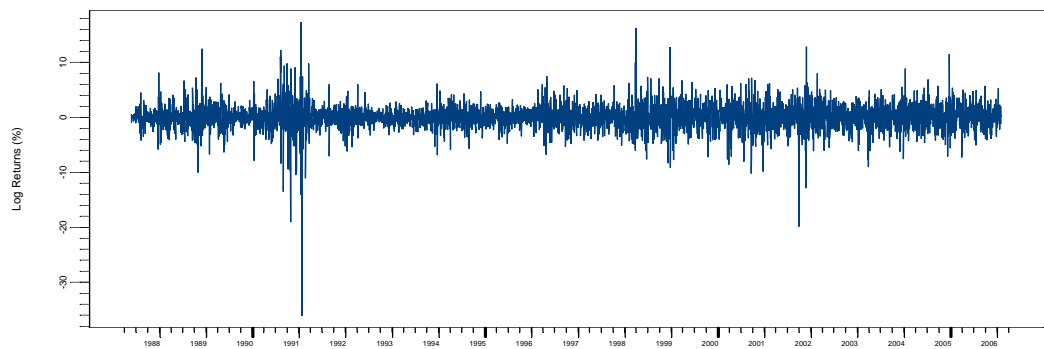


Fig. 3. Daily returns of Brent crude

Table 1
Daily returns of Brent crude, summary statistics

Mean (%)	0.0256
Std. Dev (%)	2.333
Min (%)	-37.12
Max (%)	17.33
Skewness	-0.8967
Kurtosis	20.74
Jarque-Bera	63690.41*(0.000)
Ljung-Box	50.92*(0.0002)

Table 2
AR(1)-GARCH(1,1) Estimation result

Parameter	Estimates	SE	t-value	p-value
Constant	-0.04131	0.025424	-1.625	5.214 10-2
AR(1)	0.06359	0.015376	4.136	1.798 10-5
Constant	0.07577	0.009757	7.766	4.88510-15
ARCH(1)	0.09509	0.004899	19.409	0.000
GARCH(1)	0.89502	0.005547	161.346	0.000

Table 3

This table reports the results of testing ARCH effects (LM-Test), autocorrelation (Box-Ljung) and Normality (Jarque-Bera) for the raw data as well as for the standardized residuals.

Statistic	Jarque-Bera	Ljung-Box	LM-test
Returns series	60874*(0.0000)	52.7575*(0.0001)	91.551*(0.0000)
Residuals series	512.9583*(0.0000)	12.9006(0.8816)	14.2331(0.8185)

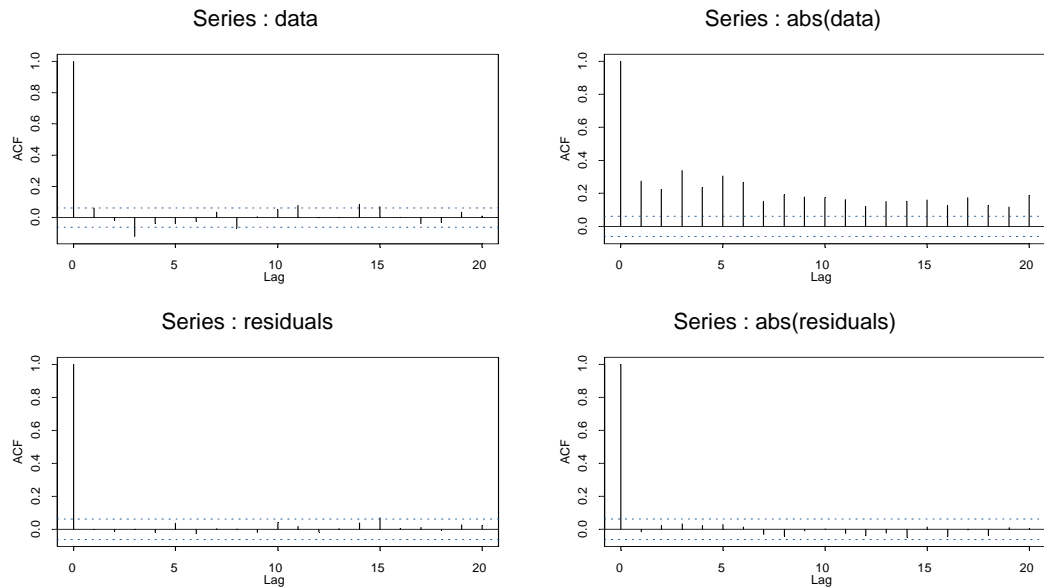


Fig. 4. Correlograms for the raw data and their absolute values as well as for the residuals and absolute residuals: iid assumption may be plausible for residuals.

Table 4

Backtesting result: Theoretically expected number of violations and number of violations obtained using: an Unconditional Normal distribution, the Historical Simulation approach, the Filtered Historical Simulation approach, the GPD model, a GARCH-model with normally distributed innovations, a GARCH-model with student t-innovations and a conditional GPD. Note that these numbers should be as close as possible to the theoretically expected one in the first line.

VaR _q	VaR.95	VaR.99	VaR.995	VaR.999
Expected	190	38	19	4
Normal	174	57	38	21
HS	161	34	23	8
FHS	192	36	18	9
GPD	185	33	18	7
Cond. Normal	188	58	40	19
Cond.Student	207	41	23	2
Cond.GPD	192	37	18	9

Table 5

Unconditional Coverage: This table reports the p-values of the unconditional coverage test. The models are successively an Unconditional Normal distribution, the Historical Simulation approach, the Filtered Historical Simulation approach, the GPD model, a GARCH-model with normally distributed innovations, a GARCH-model with student t-innovations and a conditional GPD. Note that a P-value greater than 5% indicates that the forecasting ability of the corresponding VaR model is adequate.

VaR _q	VaR.95	VaR.99	VaR.995	VaR.999
Normal	0.215	0.004	0.010	0.000
HS	0.025	0.498	0.379	0.062
FHS	0.908	0.731	0.809	0.062
GPD	0.677	0.396	0.809	0.143
Cond.Normal	0.855	0.003	0.000	0.000
Cond.Student	0.225	0.640	0.379	0.307
Cond.GPD	0.908	0.858	0.809	0.024

Table 6

Conditional Coverage: This table reports the p-values of the conditional coverage test. The models are successively an Unconditional Normal distribution, the Historical Simulation approach, the Filtered Historical Simulation approach, the GPD model, a GARCH-model with normally distributed innovations, a GARCH-model with student t-innovations and a conditional GPD. Note that a P-value greater than 5% indicates that the forecasting ability of the corresponding VaR model is adequate.

VaR _q	VaR.95	VaR.99	VaR.995	VaR.999
Normal	0.000	0.000	0.002	0.000
HS	0.072	0.477	0.216	0.171
FHS	0.903	0.613	0.892	0.171
GPD	0.075	0.401	0.197	0.339
Cond. Normal	0.832	0.006	0.000	0.000
Cond.Student	0.337	0.686	0.591	0.593
Cond.GPD	0.745	0.685	0.892	0.076

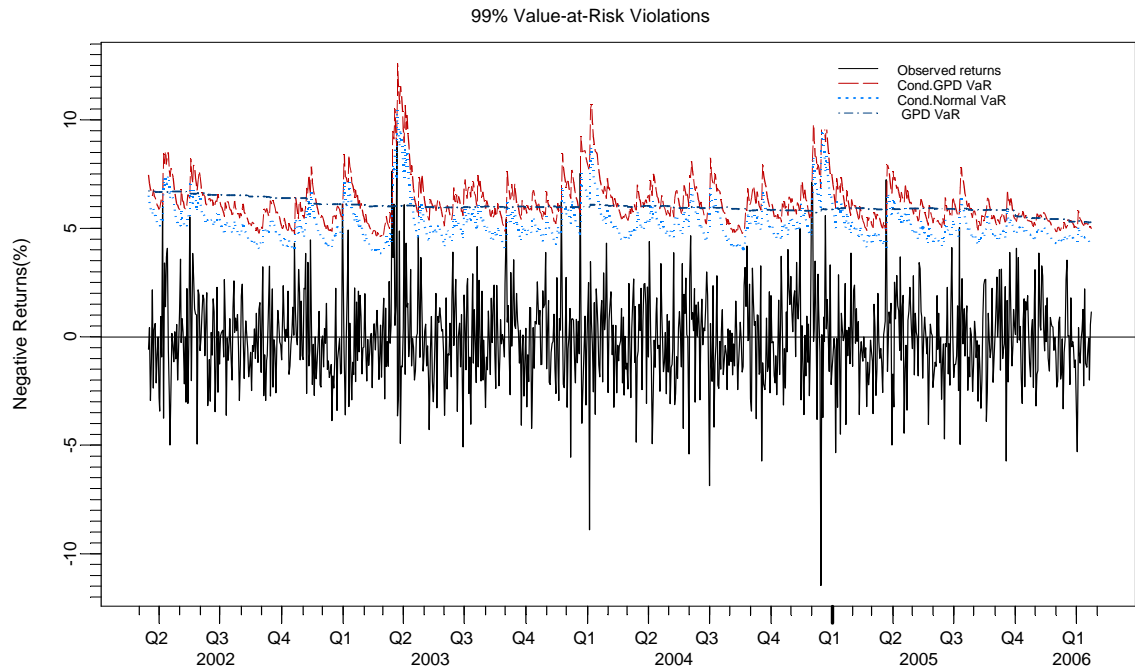


Fig.5. 1000 days of the oil returns Backtest, showing the 99% VaR estimates of conditional GPD (long dashed line), conditional normal (dotted line) and GPD (dashed line) superimposed on the negative returns. Conditional normal like conditional GPD responds quickly to the volatility dynamic, while it is all the time the less than the conditional GPD. However, the unconditional GPD fails to reacts face to a high volatility.

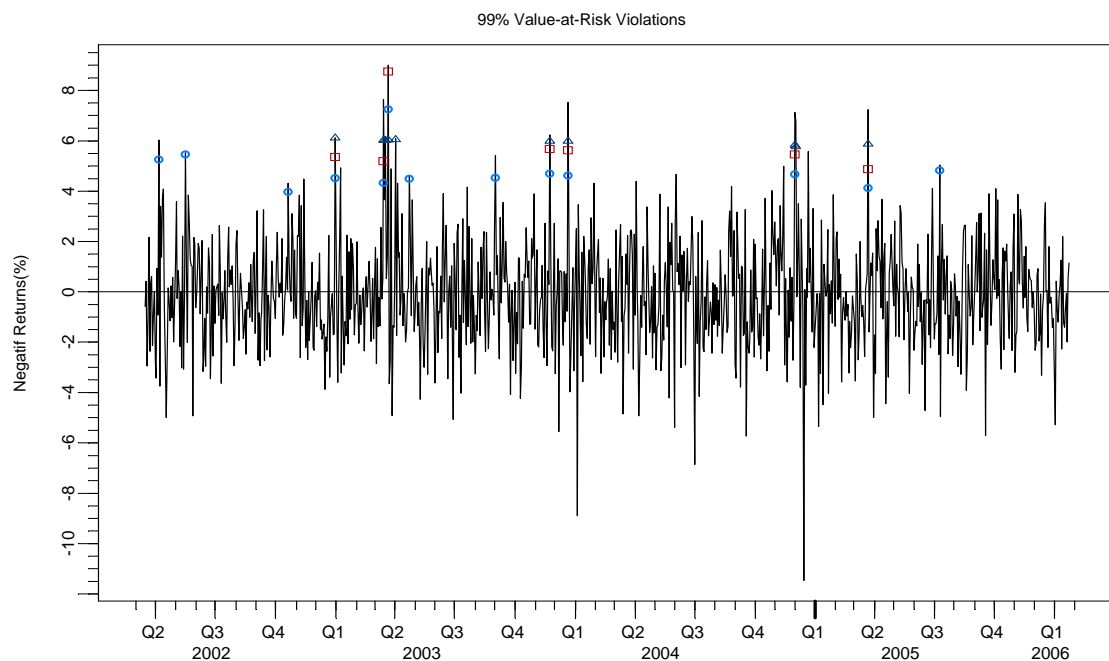


Fig. 6. Violations of 99% VaR estimates corresponding to the backtest in Figure 6. Squares, cercles and triangles denote violations of the conditional GPD, the conditional normal and the GPD respectively. The conditional normal estimate responds to volatility dynamics but tends to be violated rather more often as it fails to describe the extreme tails.

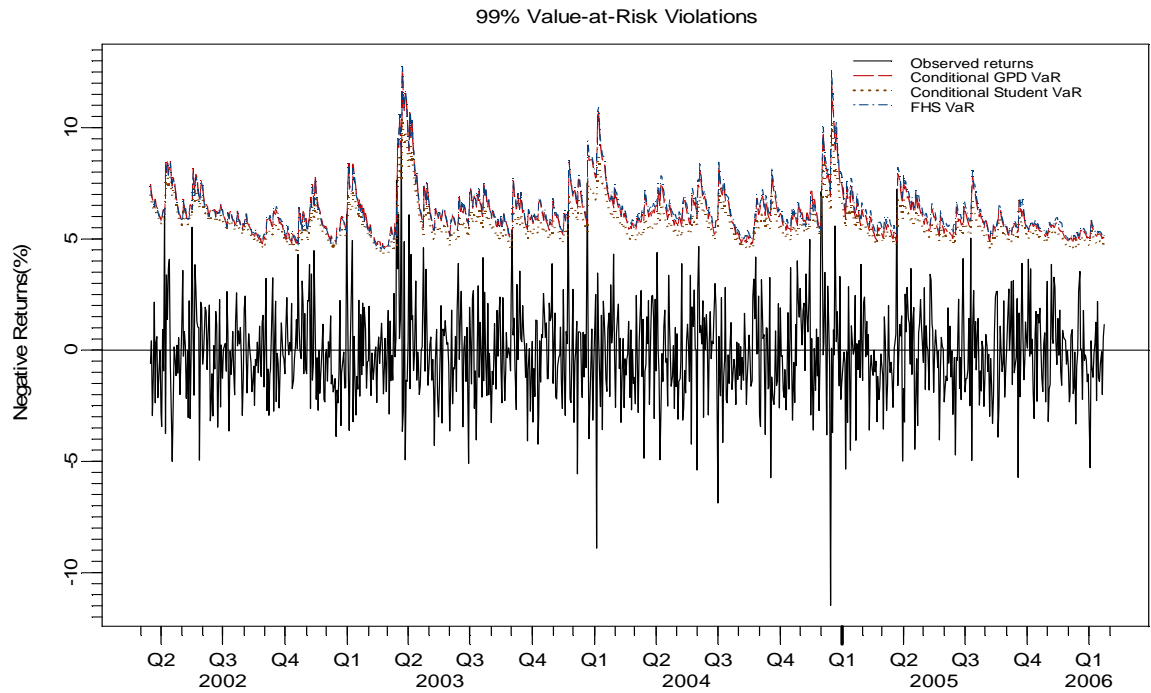


Fig. 7. 1000 days of the oil returns Backtest, showing the 99% VaR estimates of conditional GPD (long dashed line), conditional student (dotted line) and filtered historical Simulation approach (dashed line) superimposed on the negative returns. All models respond quickly the volatility dynamic, and there is no particular reason to prefer one model to another.

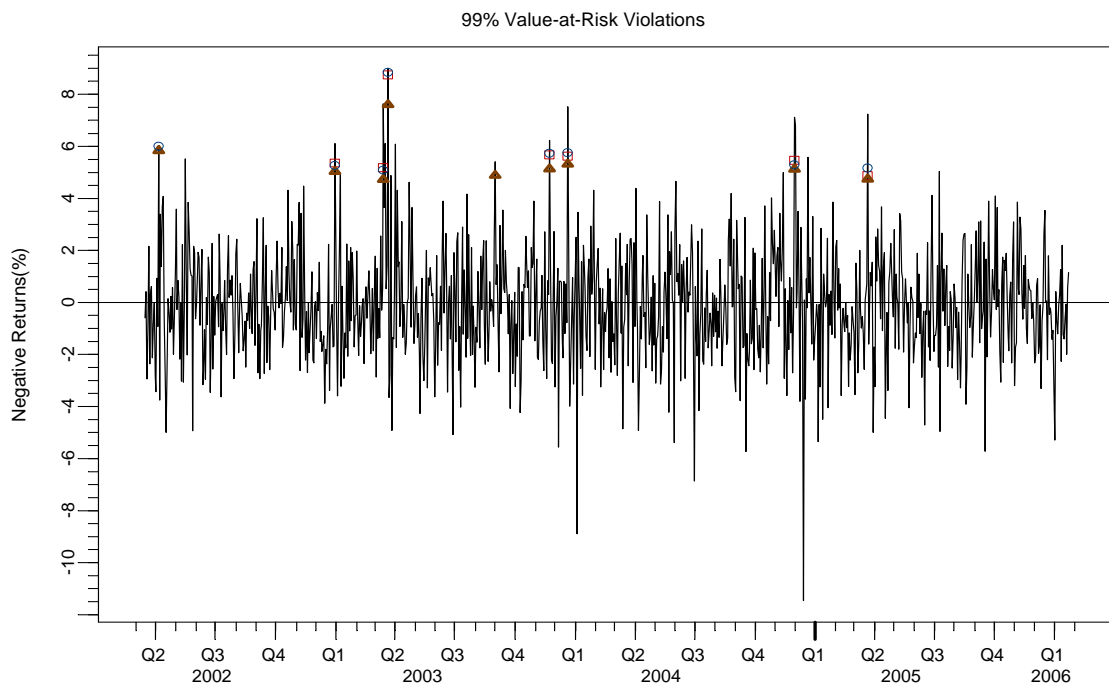


Fig. 8. Violations of 99% VaR estimates corresponding to the backtest in Figure 8. Squares, triangles and cercles denote violations of the conditional GPD, conditional student and the Filtered Historical Simulation VaR estimates respectively. There is not a particular difference between the three models; the violations points are more or less the same.